DOCUMENT RESUME

ED 308 203 TM 013 411

AUTHOR Bolton, David L.; And Others
TITLE Validity of Translations of a Cosmetology Licensure Examination.
PUB DATE Mar 89
NOTE 12p.; Paper presented at the Annual Meeting of the American Educational Research Association (San Francisco, CA, March 27-31, 1989).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Construct Validity; *Cosmetology; *English; Factor Analysis; Factor Structure; Goodness of Fit; *Licensing Examinations (Professions); Multiple Choice Tests; *Spanish; Spanish Speaking; Test Bias; Test Construction; Test Format; *Test Validity; Translation; Vocational Education
IDENTIFIERS *Florida Cosmetology Licensure Examination

ABSTRACT

A study was conducted to assess the validity of translations of two different forms of a licensing examination for cosmetologists in Florida to ensure that both Spanish and English candidates have equal chances of being licensed. The LISREL computer program was used to test the equivalence of factor structure, units of measurement, and standard errors of measurement between the original examination and the translations of the two forms. Data was from the administration of the Florida Cosmetology Licensure Examination in 1987, with 1,081 subjects taking English Form 1, 1,063 taking English Form 2, 151 taking Spanish Form 1, and 132 taking Spanish Form 2. All tests were written multiple-choice examinations containing 100 items. Initial results indicated that the goodness of fit of the translation of Form 2 to the model was not adequate. A team of consultants and Spanish cosmetologists revised this translation for further testing on samples of 1,136 examinees taking the English and 180 examinees taking the Spanish versions. Results indicate that the use of factor analysis was a feasible way to assess the construct validity of the examination. Preliminary analysis indicated a weakness in translation of one form, but analysis of data from a subsequent revision of the translation produced a more acceptable fit. At subsequent analysis, parameter modification indices indicated areas of concern, and it was concluded that further research should be done to indicate how factor analysis can help rectify areas of concern. Six tables of correlation matrices are appended. (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED308203

Validity of Translations of a Cosmetology Licensure Examination

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

DAVID L. BOLTON

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

David L. Bolton, Florida Department of Professional Regulation

Eunice E. Loewe, Florida Department of Professional Regulation

F.J. King, Florida State University

Paper to be presented at the annual meeting of the American Education Research Association

San Francisco, CA

March, 1989

TM013411

BEST COPY AVAILABLE

Objective

Because of the growing Hispanic population in Florida, the Department of Professional Regulation of the state of Florida has been required to translate licensure examinations into Spanish to accommodate this population. To ensure that both Spanish and English candidates have equal chances of being licensed, a study was conducted to assess the validity of the translations.

Theoretical Framework

Assessing the validity of translations has traditionally been a difficult task. Before the advent of more powerful statistical methods, one commonly used method of assessing translations was to have the translated document translated back into the original language to see how close the back-translated version came to the original. This method is not adequate for assessing adequacy of translations however, since skillful translators may be able to guess the content of original form from inferior translations.

One method of assessing the validity of an examination translation is to give the examination and its translation to two corresponding populations and compare the results. Comparisons across groups is difficult though, because of differences in populations tested. Maximum likelihood factor analysis provides a method of assessing the validity of an examination translation across populations. It permits comparisons of multivariate factor structure across subpopulations. It also permits comparisons of unit of measurement and the standard error of measurement.

Methodology

LISREL was used to test the equivalence of factor structure, units of measurement, and standard errors of measurement between the original examination and the translation for two different forms of an examination. Equivalence of factor structure was tested by keeping the factor loading pattern the same for the populations. Equivalence of unit of measurements was tested by keeping the factor loading values as well as the factor loading pattern the same for the populations. Equivalence of errors of measurement was tested by additionally keeping the variance of the error terms the same across all populations. Thus, the analysis provides an increasingly stronger test of the equivalence of the original and the translation.

A chi-square test is commonly used to test the fit of a model. However, the chi-square test is sample-size dependent. Other indicators of the fit of the model are the Goodness-of-Fit Indicator (GFI) and the Adjusted Goodness-of-Fit Index (AGFI). The GFI and the AGFI are indicators of the relative amount of variance and covariance explained by the model. These statistics are not dependent upon sample size. Although their statistical distribution is unknown, and therefore no test for significance

can be made, they still provide an indication of the fit of the model. Values for the GFI and the AGFI generally fall between zero and one.

Data Source

The data for the study was from the administration of the Florida Cosmetology Licensure Examination for the first six months of 1987. For each administration, both the English and Spanish forms of the examination were administered. There were two different forms of the examination, with each form being given in three consecutive months. During this time period, 1081 subjects took English Form 1, 1063 subjects took English Form 2, 151 subjects took Spanish Form 1, and 132 took Spanish Form 2.

The examination was a written clinical competency examination consisting of 100 multiple choice items which tested the candidate's ability to analyze information and make decisions about procedures and methods to use with clients. There were eight parts to the examination: Hair Coloring, Chemical Relaxing, Chemical Waving, Hair Bleaching, Hair Shaping, Thermal Styling, Manicuring, and Hair Removal. Each part was treated as an independent observed variable. It was hypothesized that all parts of the examination are measuring one common factor.

Results

Correlation matrices were calculated for both forms of the examination for both English and Spanish versions using data obtained from the data. These correlation matrices were used in subsequent analyses. The matrices can be found in Appendix A.

Initially, each of the four forms was tested using the same factor model. The chi-square values, corresponding probabilities, the Goodness-of-Fit Indices and the Adjusted Goodness-of-Fit indices for each of these four tests can be found in Table 1.

Table 1

Equivalence of Factor Structure Statistics

	Fm.1-Eng.	Fm.1-Spn.	Fm.2-Eng.	Fm.2-Spn.
Chi-square	36.80	25.61	56.15	13.60
Probability	0.012	0.179	0.000	0.850
GFI	0.992	0.960	0.987	0.976
AGFI	0.985	0.928	0.976	0.958

For the English forms, the chi-square values were significant, while for the Spanish translations the chi-square values were not significant. As stated before, the chi-probabilities are sample size dependent; the probabilities reflect the large number of English candidates and the small number of Spanish candidates. Looking at the Goodness-of-Fit and Adjusted Goodness-of-Fit

Indices, however, all indicated a relative goodness-of-fit of the model to the data.

Given the reasonable fit of the model, it was decided to test the equivalence of the units of measurement of the forms and their translations. Each form was compared separately with its translation. The factor loadings which were determined for the original were used for the corresponding translation. The chi-square for each comparison, as well as the Goodness-of-Fit and Adjusted Goodness-of-Fit Indices for each separate examination can be found in Table 2.

Table 2

Equivalence of Factor Loading Statistics

	Fm.1-Eng.	Fm.1-Spn.	Fm.2-Eng.	Fm.2-Spn.
Chi-square	**	83.82	**	109.53
Probability	**	0.001	**	0.000
GFI	0.991	0.936	0.986	0.908
AGFI	**	**	**	**

** information not available

While the GFI indicated that the fit for both Form 1 and Form 2 English examinations was excellent, the Goodness of Fit indices for the Spanish translations were lower, with the index for Spanish Form 2 being lower than that of Spanish Form 1. Even though the Goodness of Fit indices are lower, they still seem within a reasonable range.

Since the fit of the translations of Form 1 and 2 seemed to be adequate, it was decided to test whether the errors of measurements were equivalent for both forms and their translations, keeping the factor structure, factor loadings, and error terms equal. The chi-square and the Goodness-of-Fit indices for both forms can be found in Table 3.

Table 3

Equivalence of Error of Measurement Statistics

	Fm.1-Eng.	Fm.1-Spn.	Fm.2-Eng.	Fm.2-Spn.
Chi-square	**	126.13	**	181.62
Probability	**	0.000	**	0.000
GFI	0.990	0.893	0.983	0.830
AGFI	**	**	**	**

** information not available

The Goodness of Fit indices indicate that the fit of Form 1 is still reasonably good. However, the fit of Form 2 has decreased significantly. This indicates that problems exist with the translation. That there were problems with the translation of Form 2 was confirmed independently from this study by a consultant who looked at the translations and indicated a need for revisions of the translation of Form 2.

As a further study, it was decided to let a team of consultants and Spanish cosmetologists revise the translation of Form 2 and to use this translation for the examinations. This data was analyzed to assess the equivalence of factor structure, units of measurement, and standard errors of measurement across the English Form 2 and its revised translation. The further analysis was based upon a sample of 1136 taking the English form and 180 taking the Spanish revised translation. The correlation matrices for the samples can be found in Appendix A.

Again, overall factor model fit was tested initially. The chi-square value, associated probability, as well as the Goodness-of-Fit and Adjusted Goodness-of-Fit Indices for both original and translation can be found in Table 4.

Table 4

Fit of Factor Model Statistics - Revised Form 2

	Fm.2-Eng(R)	Fm.2-Spn(R)
Chi-square	46.20	34.14
Probability	0.001	0.019
GFI	0.990	0.953
AGFI	0.982	0.916

The statistics seem to indicate relatively goodness of fit of the factor model to the data.

Equivalence of factor structure and factor loadings across original and translation was then tested by setting the factor loadings equal for both analyses. The chi-square value, associated probability, as well as the Goodness-of-Fit and Adjusted Goodness-of-Fit Indices for the revised translation (R), along with the original translation (O) can be found in Table 5.

Table 5

Equivalence of Factor Loadings Statistics - Revised and Original Form 2

	Fm.2-Eng(R)	Fm.2-Eng(O)	Fm.2-Spn(R)	Fm.1-Spn(O)
Chi-square	**	**	101.52	109.53
Probability	**	**	0.000	0.000
GFI	0.989	0.986	0.930	0.908
AGFI	**	**	**	**

** information not available

As would be expected, the Goodness of Fit index for the English versions stayed the same. The Goodness of Fit index for the Spanish translation increased after revision, indicating better comparability across units of measurement between English original and Spanish translation for the revised translation.

Equivalence of factor structure, factor loadings, and errors of measurement was then tested by setting both factor loadings and errors of measurement equal for the factor model. The statistics for the analyses can be found in Table 6.

Table 6

Equivalence of Error of Measurement Statistics - Revised and Original Form 2

	Fm.2-Eng(R)	Fm.2-Eng(O)	Fm.2-Spn(R)	Fm.1-Spn(O)
Chi-square	**	**	179.23	181.62
Probability	**	**	0.000	0.000
GFI	0.987	0.983	0.855	0.830
AGFI	**	**	**	**

** information not available

Again, the Goodness of Fit index for the English stayed roughly the same. The Goodness of Fit index for the Spanish translation improved, but not as substantially as had been expected.

When the translation was given to the consultants for revision, they were provided no information from the analysis. It was now decided to look at the modification indices in the analyses to see if the translation team could have been given a clue as to what areas of the examination needed to be revised. The modification indices are provided in Table 7 for the fully equivalent model.

The content areas Thermal Styling, and Hair Removal and Scalp Treatments appear to have large modification indices under factor loadings. Under errors of measurement, Hair Bleaching and Thermal Styling have large modification indices. Thus, it appears that the translation of items in three content areas, Thermal Styling, and Hair Removal and Scalp Treatments and Hair Bleaching, appear to have problems.

Also in Table 7, the modification indices for the revised version are also provided. Overall, the magnitude of the modification indices decreased. However, for the three areas of concern, there was not much change. For Thermal Styling, the magnitude of the modification index for the factor loading decreased substantially. However, although the modification index for errors of measurement decreased, it was not reduced to acceptable levels. The error of measurement modification index for Hair Bleaching increased slightly. For Hair Removal and Scalp Treatments, the modification index decreased for the factor loadings, but increased slightly for the errors of measurement.

Table 7

Modification Indices for Factor Loadings and Errors of Measurement - Original and Revised Form 2

Content Area	Factor Loadings		Errors of Measurement	
	Orig.	Rev.	Orig.	Rev.
Hair Coloring	5.001	4.311	6.134	0.746
Chemical Relaxing	4.577	0.853	0.789	4.538
Chemical Waving	0.210	0.889	1.715	8.984
Hair Bleaching	0.814	0.130	22.448	26.388
Hair Shaping	0.244	1.895	3.061	8.122
Thermal Styling	39.849	5.201	61.329	46.627
Manicure, Pedicure, Nail Extensions	1.163	0.310	3.850	0.362
Hair Removal and Scalp Treatments	19.971	11.073	2.975	4.992

An analysis was made of the number of major or minor changes made to Spanish translation of Form 2 to see if the team of consultants made changes in the same areas as were indicated by the analysis. A minor change was defined as a slight change in the wording, where a major change was anything more than one change in stem or alternatives.

Table 8

Changes Made by Consultant

Content Area	Minor Changes		Major Changes		Total Changes	
Hair Coloring	9/21	9.5%	2/21	9.5%	11/21	52.4%
Chemical Relaxing	0/10	0.0%	2/10	20.0%	2/10	20.0%
Chemical Waving	5/10	50.0%	1/10	10.0%	6/10	60.0%
Hair Bleaching	7/22	31.8%	6/22	27.3%	13/22	59.1%
Hair Shaping	1/13	7.7%	3/13	23.1%	4/13	30.8%
Thermal Styling	8/9	88.9%	0/9	0.0%	8/9	88.9%
Manicure, Pedicure, Nail Extensions	3/8	37.5%	0/8	0.0%	3/8	37.5%
Hair Removal and Scalp Treatments	2/7	28.6%	0/7	0.0%	2/7	28.6%

There does not appear to be a relationship between the area where changes made by the consultant and the areas of concern as pointed out by the statistical analysis. However, the area which had the most changes, produced the greatest change in the modification index.

Summary and Further Research

The study indicates that the use of factor analysis provides a practical tool for assessing the construct validity of the examination across populations; in this case, testing the equivalence of the responses to different translations of the same examination. The analysis provided an overall evaluation of the translation and indicated a weakness in translation for one form of the examination. Analysis of data from a subsequent revision of the translation produced a more acceptable fit.

In a further analysis, parameter modification indices indicated areas of concern. However, further research needs to be done in determining how factor analysis information can be used to rectify areas of concern. In particular, it would be interesting to provide translators with information from the analysis and see if this makes a difference. Further research should also be done to replicate the study to assess whether the methodology proves effective with other subject matter and other circumstances.

Appendix A

Correlation Matrices for Analyses

Appendix A - Correlation Matrices

Correlation Matrix
English Form 1

1.00000	.59044	.57357	.66224	.55176	.58960	.54978	.50281
.59044	1.00000	.51227	.61463	.50619	.51838	.54030	.43305
.57357	.51227	1.00000	.54846	.47106	.52715	.45371	.43799
.66224	.61463	.54846	1.00000	.54814	.57250	.55943	.48761
.55176	.50619	.47106	.54814	1.00000	.48551	.49048	.48233
.58960	.51838	.52715	.57250	.48551	1.00000	.51497	.42236
.54978	.54030	.45371	.55943	.49048	.51497	1.00000	.44541
.50281	.43305	.43799	.48761	.48233	.42236	.44541	1.00000

Correlation Matrix
Spanish Form 1

1.00000	.40636	.26033	.39918	.37123	.38219	.47024	.42438
.40636	1.00000	.32053	.45261	.25575	.29065	.43598	.29493
.26033	.32053	1.00000	.23524	.14689	.36174	.26857	.23889
.39918	.45261	.23524	1.00000	.21122	.38590	.30841	.34690
.37123	.25575	.14689	.21122	1.00000	.23788	.34794	.34828
.38219	.29065	.36174	.38590	.23788	1.00000	.31647	.33692
.47024	.43598	.26857	.30841	.34794	.31647	1.00000	.25699
.42438	.29493	.23889	.34690	.34828	.33692	.25698	1.00000

Correlation Matrix
English Form 2

1.00000	.36129	.52358	.69078	.34016	.37185	.45739	.48948
.36129	1.00000	.27926	.39614	.14853	.24785	.31721	.23996
.52358	.27926	1.00000	.48956	.32658	.31962	.38645	.41743
.69078	.39614	.48956	1.00000	.32999	.39751	.46022	.46680
.34016	.14853	.32658	.32999	1.00000	.23195	.25397	.23041
.37185	.24785	.31962	.39751	.23195	1.00000	.32234	.31774
.45739	.31721	.38645	.46022	.25397	.32234	1.00000	.39186
.48948	.23996	.41743	.46680	.23041	.31774	.39186	1.00000

Correlation Matrix
Spanish Form 2

1.00000	.37978	.44572	.53231	.31942	.43884	.51167	.26627
.37978	1.00000	.38475	.48618	.22980	.40036	.36414	.23370
.44572	.38475	1.00000	.49320	.27390	.45749	.40388	.21311
.53231	.48618	.49320	1.00000	.23903	.51497	.39258	.20556
.31942	.22980	.27390	.23903	1.00000	.17601	.24060	.05182
.43884	.40036	.45749	.51497	.17601	1.00000	.43371	.23091
.51167	.36414	.40388	.39258	.24060	.43371	1.00000	.16402
.26627	.23370	.21311	.205561	.05182	.23091	.16402	1.00000

Correlation Matrix
English Form 2 Revised

1.00000	.39295	.51121	.66622	.35801	.38131	.48170	.50075
.39295	1.00000	.36279	.45949	.21212	.33505	.31677	.31096
.51121	.36279	1.00000	.55312	.32443	.39421	.38811	.38580
.66622	.45949	.55312	1.00000	.33613	.43227	.47004	.49291
.35801	.21212	.32443	.33613	1.00000	.19066	.25399	.22471
.38131	.33505	.39421	.43227	.19066	1.00000	.31810	.34928
.48170	.31677	.38811	.47004	.25399	.31810	1.00000	.36224
.50075	.31096	.38580	.49291	.22471	.34928	.36224	1.00000

Correlation Matrix
Spanish Form 2 Revised

1.00000	.39255	.37404	.54952	.13153	.33181	.40580	.31127
.39255	1.00000	.33148	.34178	.04056	.31073	.28660	.24827
.37404	.33148	1.00000	.40187	.19869	.44899	.35697	.31362
.54952	.34178	.40187	1.00000	.15357	.21814	.38147	.37653
.13153	.04056	.19869	.15357	1.00000	.10676	.26514	.16270
.33181	.31073	.44899	.21814	.10676	1.00000	.38464	.18584
.40580	.28660	.35697	.38147	.26514	.38464	1.00000	.27253
.31127	.24827	.31362	.37653	.16270	.18584	.27253	1.00000